

REPORT DOCUMENTATION PAGE					Form Approved OMB No. 0704-0188	
<p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p> <p>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</p>						
1. REPORT DATE (DD-MM-YYYY) 31-Dec-2003		2. REPORT TYPE FINAL		3. DATES COVERED (From - To) 01-Apr-01 to 30-Sep-03		
4. TITLE AND SUBTITLE Augmented Cognition: Amplification of Attention for Better Decision				5a. CONTRACT NUMBER		
				5b. GRANT NUMBER N00014-01-1-0706		
				5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S) Misha Pavel				5d. PROJECT NUMBER		
				5e. TASK NUMBER		
				5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Oregon Health & Science University/West Campus 20000 NW Walker Road Beaverton, OR 97006-8921				8. PERFORMING ORGANIZATION REPORT NUMBER N/A		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A				10. SPONSOR/MONITOR'S ACRONYM(S) N/A		
				11. SPONSOR/MONITOR'S REPORT NUMBER(S) N/A		
12. DISTRIBUTION/AVAILABILITY STATEMENT A. Approved for public release; distribution in unlimited						
13. SUPPLEMENTARY NOTES						
14. ABSTRACT <p>The primary purpose of this project was to investigate the underlying principles for the future development of systems capable of estimating the cognitive state of an operator, the demands of the task, and the relevant environmental conditions. The cognitive state is estimated using physiological as well as behavioral measures. Examples of the physiological measures include EEG and heart rate, and the behavioral measures include performance on the task, eye movements, etc. We developed a theoretical framework for a class of models that relate attention allocation to a measure of workload, and we demonstrated our ability to apply this framework to several tasks. Subsequently, we carried out several experiments to investigate the effectiveness of this approach within a visual search paradigm. Our results suggest that significant improvements are possible, but only in situations where the performance is indeed limited by the abilities and the state of the human operator..</p>						
15. SUBJECT TERMS Augmented cognition, visual search, workload, computational models of cognition						
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 15	19a. NAME OF RESPONSIBLE PERSON Misha Pavel	
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U			19b. TELEPHONE NUMBER (Include area code) 503-748-1155	

Augmented Cognition: Amplification of Attention for Better Decision

Final Report

January 25, 2005

Grant: NASA-NCC 2-1218

PI Name: Misha Pavel,

Department: Biomedical Engineering

Institution Oregon Health and Science University

Address: 20000 NW Walker Road

City: Beaverton State: Oregon

Zip: 97006

Country: USA

PI Phone: 503-748-1155

Email: pavel@bme.ogi.edu

Telephone: (503) 748 1155

Fax: (503) 748 1406

ABSTRACT

The primary purpose of this project was to investigate the underlying principles for the future development of systems capable of estimating the cognitive state of an operator, the demands of the task, and the relevant environmental conditions. The cognitive state is estimated using physiological as well as behavioral measures. Examples of the physiological measures include EEG and heart rate, and the behavioral measures include performance on the task, eye movements, etc. We developed a theoretical framework for a class of models that relate attention allocation to a measure of workload, and we demonstrated our ability to apply this framework to several tasks. Subsequently, we carried out several experiments to investigate the effectiveness of this approach within a visual search paradigm. Our results suggest that significant improvements are possible, but only in situations where the performance is indeed limited by the abilities and the state of the human operator.

Table of Contents

1.	Executive Summary	3
2.	Introduction.....	4
3.	Framework for Workload Modeling.....	5
3.1.	Two-Task Example	6
3.2.	Multi-Task Framework.....	8
3.3.	Multidimensional Workload Framework.....	10
4.	Empirical Aspects	12
4.1.	Data Acquisition.....	12
4.2.	Approaches to Parameter Estimation.....	12
4.3.	Multidimensional Representation	13
5.	Significance and Expected Impact	14
6.	References.....	14
7.	Appendix.....	14

1.

1. EXECUTIVE SUMMARY

The main scientific goal of this project was to carry out initial investigations of techniques for augmenting cognitive capabilities. The specific focus of this effort was to develop principles that would permit engineering developments and designs of devices and systems that would greatly enhance performance in systems currently limited by an operator's cognitive abilities. The Augmented Cognition approach consists of three components:

1. Sensing and data acquisition to support estimation of cognitive state, task demands and environmental context.
2. Cognitive state estimation using intelligent statistical pattern recognition techniques to infer concurrent resource allocation and spare capacity.
3. Artificial intelligence techniques capable of adjusting in real time the task demands to match the instantaneous processing resources of the human operator.

Because of the current capabilities in our laboratory, the focus of this project was on cognitive state estimation, the development of mathematical models, and empirical investigations of the application of the modeling efforts. The specific results of this project include:

Identification of the bottlenecks that is likely to limit performance.

Development of an initial version of a model of workload that defines a quantitative representation of workload and provides a framework for establishing a relationship between the observable biological quantities, observable performance data, and the information capacity of human operators.

Completion of a number of experiments utilizing several different implementations of cognitive support. We began to investigate the conditions that lead to significant improvements. We identified a subset of conditions under which search tasks are essentially impossible to perform without augmented cognition support.

Development of software that implements, in C++, the Warship Commander interactive game for testing a variety of interventions in support of augmented cognition systems. This software allows the designers to compute the trajectories of the individual game elements in real time, and to modify the display dynamically in response to the operators' actions and external events. The real-time capabilities are essential in order to provide a sufficiently tight feedback loop that will allow the performers to manipulate the cognitive load in real time.

2. INTRODUCTION

In order for cognitive aids to be effective performance amplifiers, they must be sensitive to the task context, the specific task demands, the environmental context, and the cognitive state of the operators. For example, an augmented cognition system may need to manipulate the amount of information flow to the operator to match his instantaneous needs and processing capabilities. Information overload as well as information deficit would likely result in suboptimal performance. The information processing capabilities depend on the instantaneous human information processing resources, and thus, are related to the instantaneous cognitive state and workload. The development of cognitive amplifiers, therefore, requires a continuous assessment of the cognitive state and understanding of the relationships among various cognitive resources and the cognitive state. An important determinant of cognitive state is a combination of skills, capabilities, fatigue, stress and workload. These determinants of cognitive state have different time courses ranging from years (e.g., training) to milliseconds. Effective cognitive amplification devices will need to be tailored to the cognitive state based on all these characteristics, but the most challenging adaptation is the one that changes most rapidly – the cognitive workload. For the purpose of this paper, we assume that workload is the critical aspect that determines the properties of cognitive amplification.

Given this assumption, there are two key questions underlying the development augmented cognition systems: (1) How to quantitatively assess the individual's cognitive state in terms of specific measures (gauges) and (2) how to use these theses measures to control the augmented cognitive aids. The only way to answer these questions is to develop explicit quantitative models that relate information processing capabilities to workload. We assert that some model is assumed implicitly whenever the empirical measurements are made and interpreted. We therefore propose to make the assumptions underlying the models explicit and directly available for the designers to construct specific engineering models.

The design of the cognitive amplifiers will, therefore, need a theoretical and a practical connection between workload assessment, performance on each task, and the control of cognitive amplifier. Such a theory has to comprise at least two components: the notion of workload and a representation of the aspects to be controlled. In the following paragraphs we describe briefly a set of basic assumptions that will enable us to develop more rigorous definitions subsequent sections.

We assume that for the execution of any particular task the human cognitive processor requires processing resources, some of which can be allocated among concurrent tasks. The key assumption is that workload is related to the available resources, and in particular, that it is monotonically decreasing with increasing available resources. Consequently, the more of the available resources are allocated to various tasks, the higher is the workload.

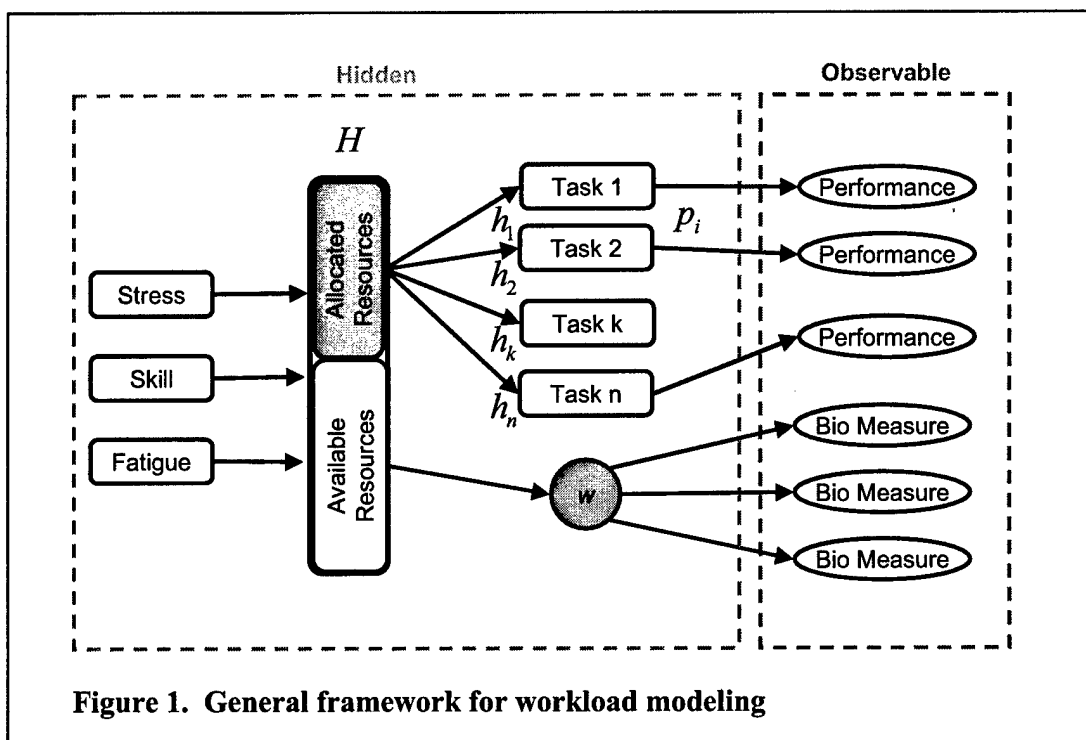
For the concept of workload to be used as a critical entity to support cognitive amplification, it is necessary to characterize the relationship between workload and the set of concurrent, individual tasks. This characterization must include (1) the contribution

of the individual task to the workload, i.e., “task load”, (2) relationship between performance on the task and the resources allocated to the task, and (3) the relationship between the difficulty, i.e. the amount of information that needs to be processed, and the resources allocated to the task.

Our approach to the development of a theoretical framework for workload will be based on an assumption that all resources are exchangeable and can be allocated to any task. We start by examining a situation with two such tasks and develop an initial notion of workload. We extend this framework to multiple concurrent tasks and develop a one-dimensional notion of workload. In the last section we discuss a modification of this approach to account for situations where some resources are bound to a class of tasks and not all resources can be allocated to arbitrary tasks.

3. FRAMEWORK FOR WORKLOAD MODELING

As we noted above, the motivation to measure workload is rooted in the relationship between workload and the available information processing resources. This notion is illustrated in terms of a diagram of a general framework in Figure 1. The diagram



identifies two components of the framework: the observable component and the unobservable constructs. The key component of the unobservable model – the cognitive resources – is represented by the block identified by H consists of two parts (1) allocated resources and (2) available resources. The allocated resources are divided among several

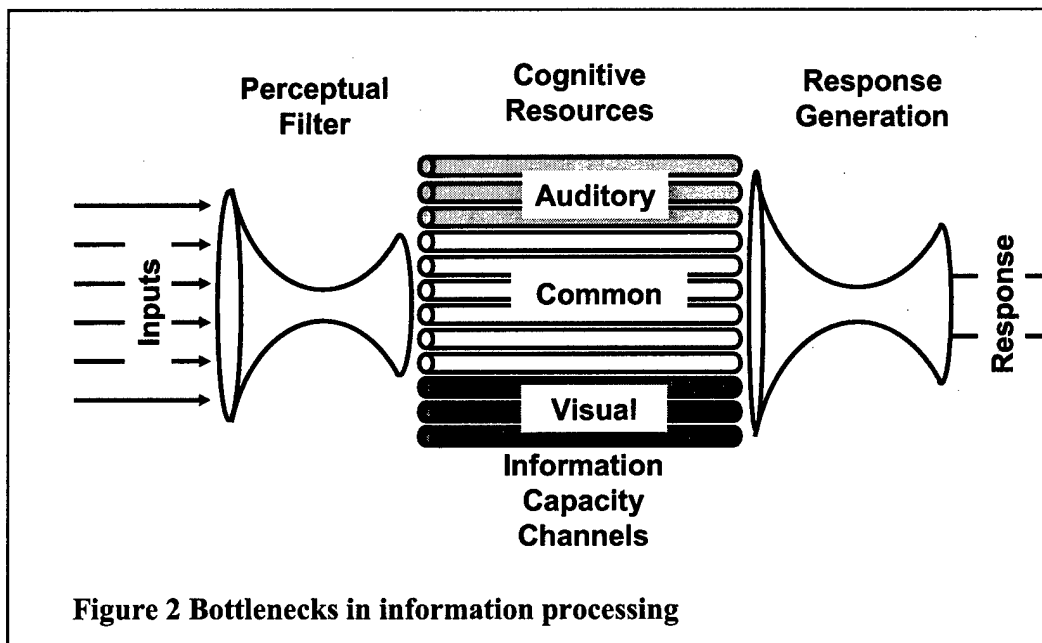
tasks, such that the resource corresponding to the i -th task is designated by h_i . The workload, denoted by w , is a variable that reflects available resources. One of the key ideas underlying cognitive amplification is to estimate the instantaneous amount of available cognitive processing resources, despite the fact that these are not directly observable. We therefore require a model that would provide a quantitative coupling between observable and unobservable aspects of this framework.

The aspects that are, at least in theory, directly observable are those that can be sensed or derived from sensory data. These are the various aspects of performance and signals from biological sensors. Although some aspects of the performance could potentially be assessed using perceptual interfaces, for the time being we would like to assess the available resources from biological data because the assessment of performance from behavioral data requires inferences that are beyond the current capabilities of the artificial intelligence methodology.

Although the ultimate goal is to focus on biological data, the starting point of the modeling effort is focused on performance. The reason for this strategy is that for the purpose of cognitive amplification of performance, it will be necessary to control the information load, the difficulty of a task as a function of the estimated available resources.

3.1. Two-Task Example

To describe the approach we start by analyzing a situation with only two tasks that



compete for the resources. This type of situation, illustrated in Figure 2, has been examined experimentally and theoretically in a number of psychological studies of dual

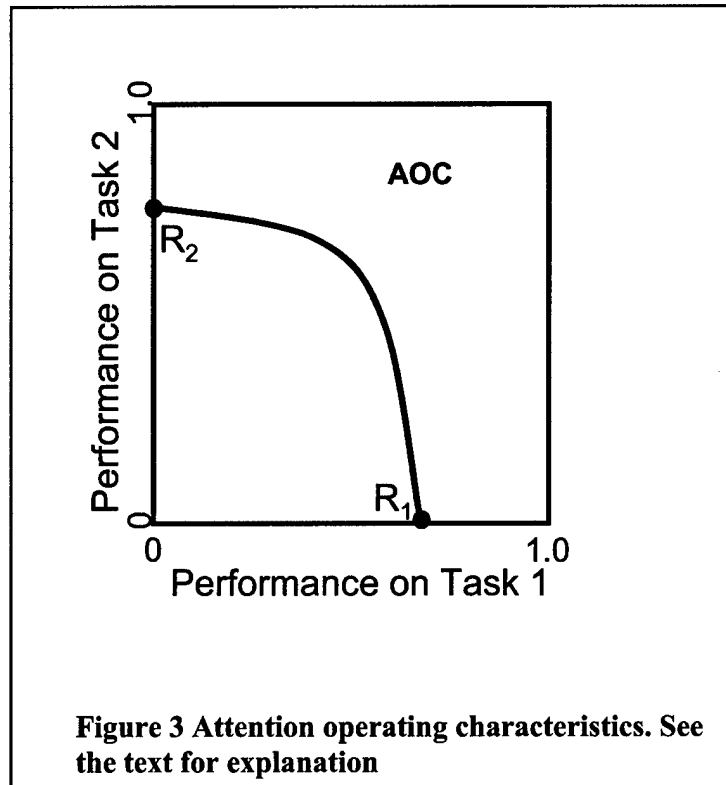


Figure 3 Attention operating characteristics. See the text for explanation

tasks. An in-depth discussion of this approach can be found in Sperling (1984), Sperling and Doshier (1986).

For the purpose of this explanation, it is useful to think of the resources as time allocated to a task. The total resources in this example is then the time allocated to the two tasks. Let's suppose that task 1 is allocated h_1 of the total H resources (amount of time), thus providing the "workload"; task 2 is the control task to be managed with respect to its task-load.

The performance on each task would presumably depend on the amount of resources allocated to each task with the resulting tradeoff between the performances on the tasks. One way to characterize these tradeoffs between two tasks performed simultaneously is the so-called attention operating characteristics (AOC) depicted by the graph in Figure 2. In this graphs the points R_1 and R_2 represent the best possible performance on each task individually. The solid curve on this graph represents the tradeoff between the performance on the two tasks and the particular shape of the curve reflects the degree to which the two tasks share resources.

The parameter of the AOC is the amount of resources allocated to task 1, h_1 . The fundamental relationship that underlies the AOC is therefore the dependence of performance on each task on the amount of resources allocated to this task. Let's denote this performance by

$$p_i = g_i(h_i) \quad (1)$$

where p_i is the performance on the i -th task and g_i is a monotonic function. If the total amount of resources is constant, and h_1 is allocated to task 1, then the performance on task 2 will be given by

$$p_2 = g_2(H - h_1) \quad (2)$$

These two relationships then specify completely the shape of the AOC. At this point it is possible to consider a definition of workload that is generated by the processing resources allocated to task 1. At this point the only constraint on the definition workload is that it must increase with the decreasing available resources and decrease with the total amount of resources. These considerations suggest two simple models. First is based on the assumption that resources are represented by an interval scale and has the form

$$w = h_1 - H = g_1^{-1}(p_1) - H. \quad (3)$$

The second model is based on the assumption of a ratio scale and has the form

$$w = \frac{h_1}{H} = \frac{g_1^{-1}(p_1)}{H}. \quad (4)$$

Both of these models have the desired properties and are in fact are closely related to each other by a logarithmic transformation.

3.2. Multi-Task Framework

To extend the definition of workload from Section 3.1 to multiple tasks requires an extension of the AOC representation to multiple dimensions. To describe a multidimensional AOC surface we start with some definitions and assumptions. First we define a performance vector $\vec{p} = [p_1, p_2, \dots, p_n]$ where each element corresponds to the performance on the corresponding task. The unobservable resource allocation to each task is described by a vector $\vec{h} = [h_1, h_2, \dots, h_n]$. Given these definitions, the workload model can be specified by the following assumptions:

1. Total cognitive resources available for an individual depends on his skill level, fatigue and stress $H(S, K, F)$, where S is stress, K is skill level and F is fatigue.
2. Workload is a function of available and allocated cognitive processing resources, $w = (h_1, h_2, \dots, h_n, H)$.
3. The resources allocated to different tasks are combined additively and the sum cannot exceed the total available resources,

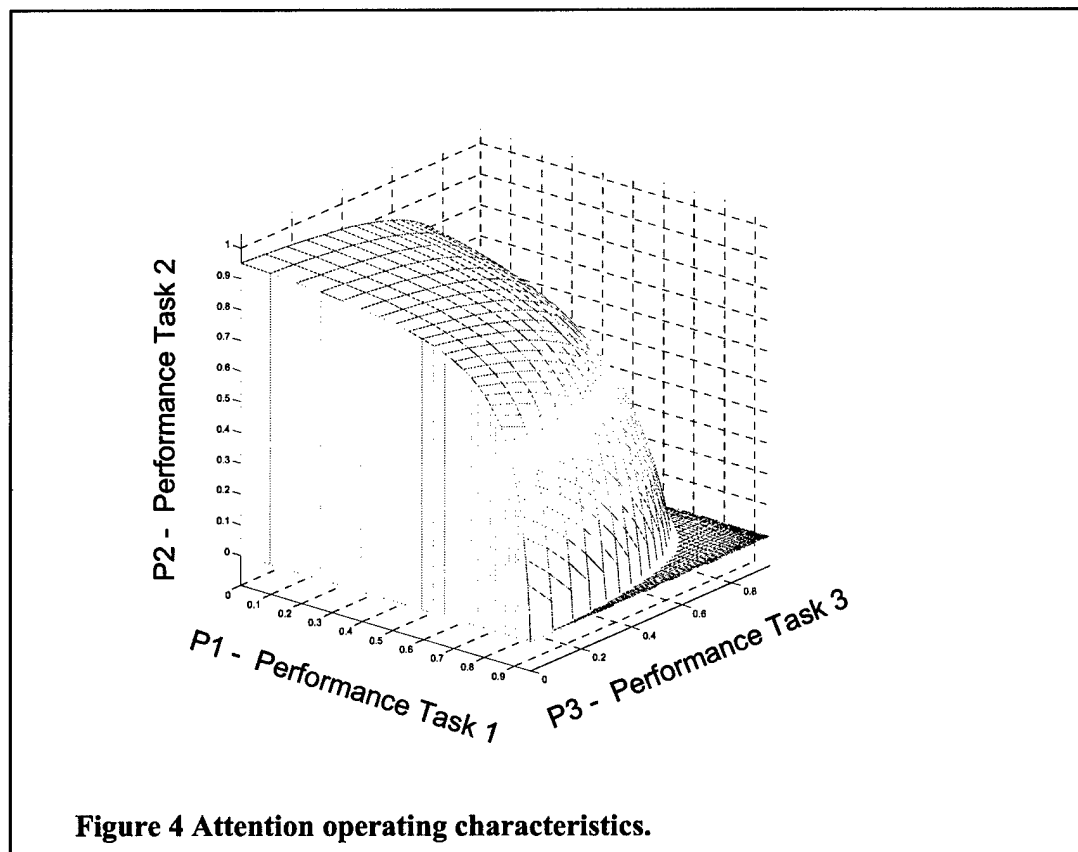
$$\sum_i h_i \leq H \quad (5)$$

4. Task performance on a particular task depends only on the amount of resources allocated to the task and the relationship between the performance on the i -th task is given by a monotonic performance function g in a similar manner to that in Equation (1)

$$p_i = g(h_i). \quad (6)$$

5. Task performance is conditionally independent, given resource allocation.
6. Biological measures of workload depend only on the specific set of concurrent tasks with a possible parameter representing free resources..

Although the total amount of resources that depend on the operator's knowledge,



fatigue, and stress, these quantities change relatively slowly in comparison to the dynamics of workload, and we therefore drop the reference to these variables, and denote the resources by H .

Perhaps the most critical notion is the additivity of resources given by assumption 3. This notion is consistent with the intuition arising from allocating physical resources such as time, energy, money, amount of equipment, or processing cycles.

The goal of this effort is to develop a parsimonious way to use biological measurements to estimate the resources that are available to the human operator for a given task. A useful class of representations of workload consistent with the above assumptions is given by the so-called difference models that have the form:

$$w(h_1, h_2, \dots, h_n, H) = f \left[u \left(\sum_i h_i \right) - u(H) \right], \quad (7)$$

where the functions f and u are monotonic functions. The simplest form of the difference model is obtained with u being an identity function with the result.

$$w(h_1, h_2, \dots, h_n, H) = f \left(\sum_i h_i - H \right). \quad (8)$$

With this model, the value of w would be determined by biological measurements. To estimate the available resources for a specific task would require only the measurement of w and compute the inverse of the empirically determined function f .

An alternate useful model is obtained by assuming that u is a logarithmic function. In that case we can write

$$w(h_1, h_2, \dots, h_n, H) = f \left[\log \left(\sum_i h_i \right) - \log(H) \right] = f' \left[\frac{\sum_i h_i}{H} \right], \quad (9)$$

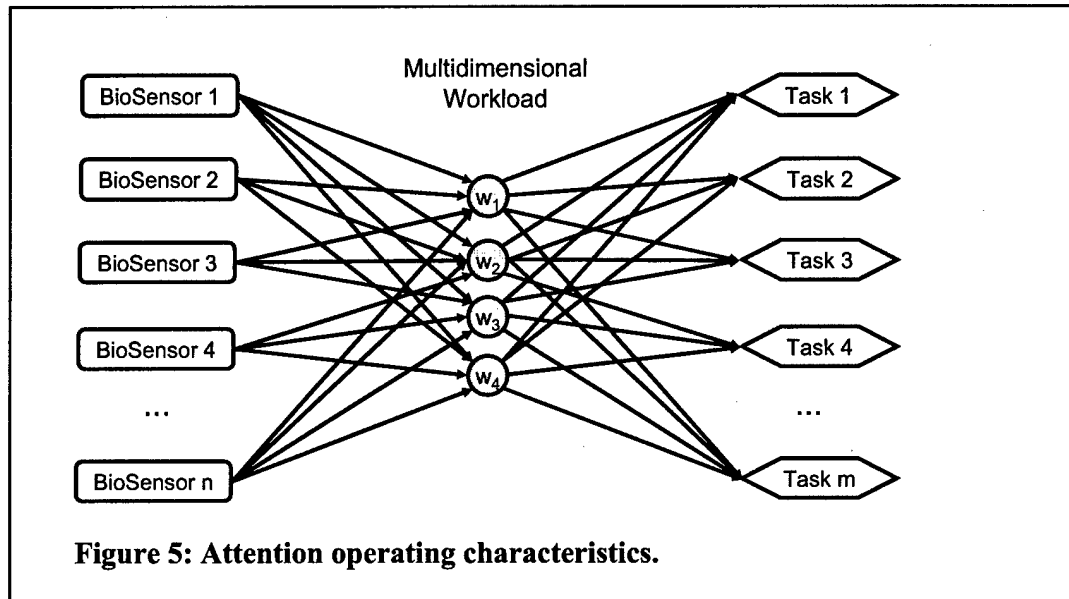
where f' is a monotonic function. This model has the homogeneity property that maintains the workload constant if both the allocation and the total resources are changed by the same factor.

It is possible that the simple formulation in Equation (8) will not be consistent with empirical data. We will discuss the possible modifications in subsequent sections

3.3. Multidimensional Workload Framework

The final step in the generalization of the workload model is based on the notion that the parallelism of brain mechanisms and the concomitant cognitive processes are likely to share a portion processing resources, but that there are resources that are dedicated to specific domains. This notion was implicitly illustrated in Figure 1 by indicating the resources dedicated to each of the perceptual channels. The empirical data consistent with this notion arise from

experiments where, for example, the detection of visual flash does not interfere with the detection of a tone, regardless of the instructions to the subject. This notion led us to postulate that the notion of workload must be extended to a



multidimensional representation in terms of a vector, as shown in Figure 5. The dimensionality of the workload vector must be sufficiently large to embody any potential conditional independence among the measures of performance on the different tasks. This means that the performance on any particular task will depend on one or more of the elements of the workload.

The following are examples of cognitive processes that represent important dimensions of cognition underlying cognitive performance and quality decision-making:

- Large capacity of accurate working memory
- Integration of multiple information sources at different levels of abstraction
- Logical reasoning in complex and abstract problem spaces
- Enumerating all possible outcomes
- Incorporating and combining uncertainties and utilities
- Predicting consequences and discounting future values appropriately
- Generating sufficiently randomized actions (in competitive games or military engagements)
- Appropriate distribution rather than focused attention – for anticipation of potential barriers during problem formulation
- Focusing attention on the most critical issues during decision making

4. EMPIRICAL ASPECTS

A complete investigation of the suitability and limitations of this modeling approach was well beyond the scope and the resources allocated to the present project. We, therefore, used our empirical efforts to illustrate the approach. An application of this approach to search is presented in an article attached as an Appendix. In the remainder of this section we will focus on the description of the possible data sources and then on an estimate of the dimensionality of workload derived by analysis of the performance on a variety of simultaneous tasks.

4.1. Data Acquisition

The sensor and data acquisition system capable of supporting an augmented cognition system must be capable of estimating cognitive state, task demands and environmental context. Consequently, the data would include physiological, behavioral and environments information. Physiological data useful for this purpose include electroencephalography (EEG), near infrared (NIR), heart rate, plethysmography, galvanic skin response, etc. Behavioral data include position, movements (specifically eye movements), vocal emissions, etc. Environmental sensors include microphones, thermometer, video cameras, infrared cameras, etc. In addition, the environmental information includes a database that is as complete as possible in terms of the terrain and situation information obtained from other INT sources.

4.2. Approaches to Parameter Estimation

For simplicity we first assume the one-dimensional representation of workload. This is possible in situations with only two tasks and the workload determines the tradeoff between the two tasks. In order to estimate the available resources using the explicit quantitative model of the workload, we need to be able to estimate the various components of the model. In particular, we need to estimate the function f and the performance functions g_i . Depending on the additional simplifying assumptions that one is willing to make, there are several approaches to estimate empirically the parameters of the model.

The most principled and most costly approach is based on the classical scaling and measurement-theoretic approach motivated by the following substitution into Equation (8),

$$w(h_1, h_2, \dots, h_n, H) = f\left(\sum_i h_i - H\right) = f\left(\sum_i g_i^{-1}(p_i) - H\right) \quad (10)$$

which is permissible because g is a monotonic function. We can then use approaches such as conjoint measurement to determine individual performance functions. We can also use a particular task as a standard and evaluate each other task with respect to that standard task. These approaches have the advantages that are relatively free of

assumptions, but this flexibility is reflected by the large quantity of data required to determine these functions.

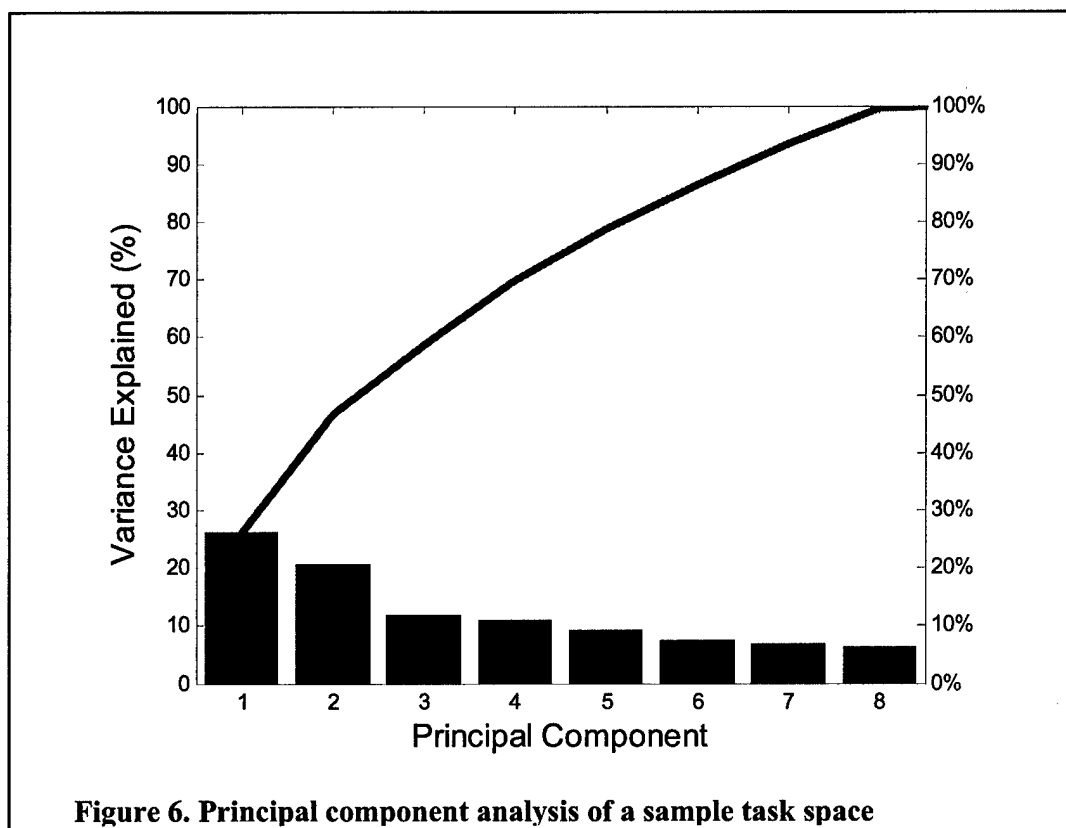
An alternative approach that is more economical and data efficient is based on the assumption of the functional form for the performance functions g . For example, a reasonable assumption is that each g belongs to a family of exponential functions,

$$g_i(h) = a(1 - e^{-\alpha_i h}) \quad (11)$$

with two parameters that could be estimated with much less data than the measurement of the entire function g . This approach was used in the paper attached as an Appendix.

4.3. Multidimensional Representation

In order to illustrate the multidimensional nature of the workload construct, we examined the results of a pilot study performed in conjunction with the early stages of our collaboration with Honeywell. In this experiment, the subject was playing a war game and his performance was evaluated on a number of dimensions. As the first attempt to estimate the dimensionality of the task space, we performed a principal component analysis. The results of this analysis are shown in Figure 6.



These results indicate that with a four-dimensional representation of the task space, the system could account for more than 70% of the variance. We should, therefore look for at least a four-dimensional representation of the sensor data (gauges). We note that these analyses are based on the assumption of linear transformation, and by allowing nonlinear transformations, a further reduction in dimensionality may be possible.

5. SIGNIFICANCE AND EXPECTED IMPACT

Although the analysis, modeling and empirical pilot test are far from an evaluation of this approach, it is obvious that a successful implementation of these techniques would be very beneficial to our national defense. One of the potential benefits of this approach is the ability to anticipate when is a person is interruptible and predict his performance on each of multiple of tasks.

6. REFERENCES

- Bertelson, P. (1998). Starting from the ventriloquist: The perception of multimodal events. In: Sabourin, M., Craik, F., et al. *Advances in psychological science, Vol. 2: Biological and cognitive aspects*, 419-439. Erlbaum.
- Koopman, B. O., (1956) The theory of search: Target detection. The Journal of the Operation Research Society of America, 4, 503-531.
- (2003) Pavel, M, Wang, G., Li, K., Kehai Li, Augmented cognition: Allocation of attention, Proceedings of 36th Hawaii International Conference on System Sciences January 6-9, 2003, Big Island, HI, USA. IEEE Computer Society, 2003, ISBN 0-7695-1874-5
- Shaw, M.L. (1978). A capacity allocation model for reaction time. Journal of Experimental Psychology: Human Perception & Performance, 4, 586-598.
- Shaw, M.L., & Shaw, P. (1977). Optimal allocation of cognitive resources to spatial locations. Journal of Experimental Psychology: Human Perception and Performance, 3, 201-211.
- Sperling, G. (1984). A unified theory of attention and signal detection. In R. Parasuraman and D. R. Davies (Hrsg.), *IVarieties of attention* (S. 103-181)
- Sperling, G. and Doshier, B.A. (1986). Strategy and optimization in human information processing. In *Handbook of Human Perception and Performance: Volume 1*, K. Boff, L. Kaufman, and J. Thomas (eds). New York: Wiley.
- Triesman, A. (1985). Preattentive Processing in Vision. *Computer Vision, Graphics, and Image Processing* 31, 156-177.

7. APPENDIX

Attached are two items. First is a paper reporting our earlier results and the second is a description of our software.

Pavel, M, Wang, G., Li, K., Kehai Li, Augmented cognition: Allocation of attention, Proceedings of 36th Hawaii International Conference on System Sciences January 6-9, 2003,